CHAPTER 4

HIGH-RESOLUTION PROCESSING TECHNIQUES FOR TEMPORAL AND SPATIAL SIGNALS

H.J. WHITEHOUSE
Naval Ocean Systems Center, San Diego (USA)

B. BOASHASH
University of Queensland (Australie)

P.M. SPEISER
Naval Ocean Systems Center, San Diego (USA)

Chapter 4

High Resolution Processing Techniques for Temporal and Spatial Signals

H. J. Whitehouse[1]
B. Boashash[2]
J. M. Speiser[3]

4.1 Introduction

The first modern generation of signal processing algorithms used Fourier methods. Their theoretical basis was provided by the convolutions which occur in time-invariant linear systems and the important role of sinusoids as eigenfunctions of every such system. A second class of signal analysis and parameter estimation methods use a general linear model similar to the one used in the statistical literature. Representative methods in this class are the linear predictive and signal subspace methods for parameter estimation. A third class of signal processing methods based on time-frequency analysis of nonstationary signals began with the work of Wigner (1932) and

[1]Naval Ocean Systems Center, San Diego, CA 92152-5000
[2]CRISSP, University of Queensland, Brisbane 4067, Australia
[3]Naval Ocean Systems Center, San Diego, CA 92152-5000

Ville (1948) and has seen extensive theoretical development and the beginning of hardware implementation in the 1978-1988 period. Recent work has combined linear predictive modelling with Wigner-Ville analysis. The near future will see the combination of time-frequency distribution (TFD) methods with both eigensystem-based signal analysis and wideband ambiguity function and wideband WVD analysis in which Doppler is treated as a time compression or expansion.

This paper provides an updated discussion of high resolution processing techniques for temporal and spatial signal processing, with particular emphasis on methods applicable to underwater acoustic signal processing. The paper consists of three parts: The first part treats direction finding and spectral estimation for stationary random processes. The second part discusses time-frequency analysis for nonstationary signals using both the Wigner-Ville distribution (WVD) and an autoregressive modified WVD. The third part discusses the application of these techniques to radar and sonar imaging.

4.2 Direction Finding

This section provides an overview of selected beamforming and direction finding (df) techniques and their computational requirements. It provides a description of algorithms and an assessment of their range of applicability. Last it recommends areas for future study. With a few noted exceptions, it considers only narrowband processing - i.e. it is assumed that the data snapshots are the Fourier coefficients in one frequency bin for each of the array elements for the currrent observation interval, and that different frequency bins are processed independently.

4.2.1 Algorithms Considered

- Classical Beamforming
- MVDR and Capon's Method
- DF by Linear Predictive Spectral Estimation
- MUSIC
- Johnson & DeGraf
- Other MUSIC variants
- Pisarenko Harmonic Retrieval

- ESPRIT
- ESPRIT variants
- Cadzow's Eigenvector Approximation Method
- Maximum Likelihood

4.2.2 Principal Notation for Direction Finding

$\mathbf{x}$	data vector
$\mathbf{x}^{(k)}$	data vector for kth snapshot
K	total number of snapshots currently available
M	number of elements in array
B	number of beams
W	total processed bandwidth
$\mathbf{a}(\mathbf{u})$	arrival phase shift vector for look direction $\mathbf{u}$. As a function of $\mathbf{u}$, also called the "array manifold"
A	1) general matrix 2) steering phase shift matrix $A = (\mathbf{a}(\mathbf{u}_1) \cdots \mathbf{a}(\mathbf{u}_D))$
D	1) number of arriving wavefronts (can be larger than the number of sources) when multipath arrivals are considered) 2) diagonal matrix in svd or eigenvalue decomposition
$\mathbf{s}$	steering vector. alternate notation for $\mathbf{a}(\mathbf{u})$
E	usually denotes statistical expectation
H	used as a superscript to denote Hermitian transpose
P	1) source correlation matrix 2) unitary matrix
Q	1) unitary matrix 2) matrix with orthonormal columns - part of a unitary matrix
U	usually used to denote upper (or right) triangular matrix. In the numerical analysis literature, R is usually used to denote a right triangular matrix. However, since the signal processing literature usually uses R for a correlation matrix, we will usually use U instead, except in the names of numerical algorithms, such as QR decomposition and the QR eigensystem method.
R	spatial correlation matrix. $R = E\mathbf{x}\mathbf{x}^H$ When it is desired to emphasize the dependence on $\mathbf{x}$, we will use $R_{\mathbf{x}}$
$R_{\mathbf{n}}$	1) noise spatial correlation matrix 2) scaled noise correlation matrix
$\mathbf{w}$	weight vector (combined amplitude and phase)

PDQ^H	factors in singular value decomposition. P is unitary D is diagonal, and Q is unitary
QDQ^H	factors in eigenvalue decomposition of a Hermitian matrix. Q is unitary, D is diagonal.
n	noise vector
f	1) vector of complex amplitudes for wavefronts arriving at the array 2) frequency
σ^2	noise power. possibly after spatially prewhitening

4.3 Overview and Algorithm Assessment

4.3.1 Classical Beamforming

We consider classical beamforming only to establish notation and terminology that will be used in discussing additional algorithms. Standard references for classical beamforming include Ma [30] and Steinberg [65]. For most algorithms, we will regard the frequency-domain output of the current data snapshot for the array as a single complex vector, $\mathbf{x}$. The data model used is:

$$\mathbf{x} = A\mathbf{f} + \mathbf{n}$$

Classical frequency domain beamforming via phase shift and sum is the frequency domain equivalent of delay and sum time domain beamforming. For each look direction u, it forms the inner product of the steering phase shift vector s = a(u) with the data vector x. This inner product, b = [a(u),x] is the spatial equivalent of a matched filter. A derivation for general array geometries with simplifications for special array geometries has been provided by Speiser, Whitehouse, and Berg [59]. The function a(u) is sometimes called the "array manifold" [52].

Shading may be used to reduce sidelobe height, at the expense of reduced resolution. Special array geometries permit the beamforming to be performed via an FFT. These include uniform line and planar arrays, circular and cylindrical arrays, and logarithmic linear and spiral arrays [59].

4.3.2 MVDR Beamforming

Minimum Variance Distortionless Response (MVDR) beamforming attempts to reject interfering sources while maintaining unit gain and zero phase shift for each look direction [40] [43].It is sometimes referred to as "Optimum Array Processing" [40]. Unlike classical shaded beamforming, it does not

try to create uniformly low sidelobes in the response pattern, but only to form nulls in the direction of interfering sources. For each steering vector $s = a(u)$, it forms $b = [w, x]$, where the weight vector w is chosen to minimize the expected power output, while providing unit gain and zero phase shift for a signal arriving from direction u. i.e. pick w to minimize $E|b|^2$ subject to $[w,s] = 1$.

$$w_{opt} = \frac{R^{-1}s}{[R^{-1}s, s]}$$

MVDR is not a high resolution method per se, but it is recommended as a baseline for comparison. It is quite robust, useful with any geometry, and very well understood, both in terms of its performance and parallel algorithms and architectures for real time implementation, including variants incorporating multiple linear constraints [54] [32].

Apart from resolution MVDR does have one limitation: coherent multipath arrivals in the same frequency bin may cause it to effectively lose sensitivity. That is, when the beam is steered to one of the arrivals, even though the weight vector provides unit gain in that direction, the response to a coherent multipath arrival may reduce the signal component of the output.

4.3.3 Capon's Method

Capon's method estimates the angular spectrum, or power versus arrival direction as $1/[R^{-1}s, s]$ [17], [18], [26]. This is the average power output of an MVDR beamformer with steering vector s. The signal processing literature (including Schmidt's papers) frequently refers to Capon's method as "maximum likelihood". This is a source of confusion, since it does not perform maximum likelihood parameter estimation, and does not have the high resolution of true maximum likelihood spectrum estimation/direction finding. We will regard Capon Maximum Likelihood as just an auxiliary output easily available from an MVDR beamformer.

4.3.4 DF by Linear Predictive Spectral Estimation

Linear predictive spectral estimation determines the weights of a prediction error filter for a wide-sense stationary random sequence.

Such a filter converts the sequence into white noise. Since the spectral density of the output of a filter is the product of the input spectral density with the magnitude squared of the filter's transfer function, the input spectral density function is proportional to the reciprocal of the magnitude squared of the filter's transfer function. When the input process has an

all-pole spectrum, the optimum prediction based on the infinite past is the same as the optimum predictor using a number of past samples equal to the number of poles. This method is sometimes called Maximum Entropy Spectral Estimation - more specifically the Yule-Walker version of Maximum Entropy Spectral Estimation [26],[35]. It has been proposed as a direction finding technique by utilizing the equivalence of spectral estimation for a wide-sense stationary random process with power versus arrival direction estimation for a uniformly spaced line array receiving signals from uncorrelated sources [31]. Linear predictive spectral estimation can provide high resolution, but require special precautions to avoid line splitting and bias errors [35]. The optimization of a signal processing function analogous to a thermodynamic function can be used not only to derive linear predictive estimators, but also can be used in a new class of estimators which can be used with a small number of data snapshots [56].

4.3.5 MUSIC

The MUSIC method is at this time the most thoroughly studied and best understood of the eigensystem based direction finding techniques. It was independently developed by R. Schmidt [52], [53] in the U. S. and by the team of L. Kopp and G. Bienvenu in France [6]. Schmidt's formulation uses the generalized eigensystem of the data correlation matrix R_X and the scaled noise alone spatial correlation matrix R_{11} . The Bienvenu and Kopp formulation (and many others who have studied MUSIC) assume that the noise is spatially white or the data has been spatially prewhitened, so that $R_{11} = \sigma^2 I$. The noise power, σ^2 is regarded as an unknown parameter. In this case, the generalized eigensystem reduces to the ordinary eigensystem of R_X alone. In discussing MUSIC it is necessary to distinguish between the behavior with perfect knowledge of the data correlation matrix and scaled noise alone correlation matrix, versus behavior with estimated or modelled correlation matrices.

MUSIC requires several assumptions:

1. The number of arriving wavefronts (in the frequency bin being analyzed) is strictly less than the number of elements in the array.
2. The columns of the arrival phase shift matrix, A, are linearly independent.
3. The source correlation matrix, P, is strictly positive definite. (sources can be coherent - but not perfectly coherent)
4. The noise spatial correlation matrix is nonsingular. (This requirement is unlikely to ever cause a problem)

With perfect knowledge of the data correlation matrix and scaled noise spatial correlation matrix, MUSIC proceeds as follows:

Solve the generalized eigensystem

$$R_{\mathbf{x}}\mathbf{z} = \lambda R_{\mathbf{n}}\mathbf{z}$$

If the generalized eigenvalues are numbered in decreasing order, the last $M - D$ are equal, and the first D generalized eigenvalues are strictly greater. The span of the generalized eigenvectors corresponding to the last $M - D$ generalized eigenvalues is called the noise subspace. Each vector in the noise subspace is orthogonal to the arrival phase shift vector $\mathbf{a}(\mathbf{u}_d)$ from each source direction. The number of sources is estimated as $D1 = M - M_0$ where M_0 is the multiplicity of the smallest generalized eigenvalue.

The arrival directions are then determined by looking for array manifold vectors $\mathbf{a}(\mathbf{u})$, orthogonal to the noise subspace. This determines the arrival phase shift matrix A. Then the source correlation matrix P, can be computed as

$$P = (A^H A)^{-1} A^H (R_{\mathbf{X}} - \lambda_{min} R_{\mathbf{n}}) A (A^H A)^{-1}.$$

Of course it is preferable to use the pseudovinverse of A, computed by an SVD, rather than forming

$$(A^H A)^{-1} A^H$$

in order to avoid problems of ill-conditioning caused by near collinearity of columns of A when signals arrive from nearly the same direction.

Early simulations by Schmidt, using perfect correlation matrices, showed that MUSIC had far higher resolution than classical beamforming and Capon's maximum likelihood method, and did not have the bias problems of the maximum entropy method.

It is important to note that the MUSIC "DOA spectrum" is simply a function that peaks in directions where the steering vector is nearly orthogonal to the noise subspace. That is, peaks correspond to arrival directions, but the DOA spectrum is not a measure of arrival power versus direction. Furthermore, the processing is highly nonlinear, so there is no simple interpretation of the apparent peak-to-sidelobe ratio.

The previous comments must be modified somewhat when estimated correlation matrices are used, but MUSIC still remains a very strong contender for direction finding. When estimated correlation matrices are used:

1. Instead of a smallest generalized eigenvalue with multiplicity (M - D), there is a spread of small generalized eigenvalues. If we pick too large a dimension for the noise subspace, then signals will be missed. If we pick somewhat too small a dimension for the noise subspace,

the method degrades gracefully: Spurious peaks will be found in the DOA spectrum, but these can be eliminated through estimation of the source correlation matrix and eliminating peaks corresponding to zero power.

2. With ideal data and scaled noise correlation matrices, MUSIC can localize in the presence of highly correlated (but not perfectly correlated) multipaths. In simulations using estimated correlation matrices, problems have been reported when the correlation between multipath arrivals was around .7 or so.

A survey of work on MUSIC and related eigenvector techniques through 1985, with detailed discussion of source covariance matrix estimation and a signal estimation techniques is available [61].

4.3.6 Johnson & DeGraaf

The method of Johnson and DeGraaf is a slight modification of MUSIC. In MUSIC, the generalized eigenvalues are only used in deciding on the dimensionality of the noise subspace, not in forming the DOA spectrum - i.e. all the basis vectors for the noise subspace are weighted equally in the DOA spectrum. Johnson and DeGraaf use a modified DOA spectrum, in which the eigenvector weighting depends on the corresponding eigenvalue, so that in effect the decision as to the dimension of the noise subspace (or equivalently the number of signals present) is a soft decision [27].

In simulations by the inventors, it was found to provide the same performance as MUSIC when the correct number of signals was chosen for both, but to be more robust when errors were made in picking the number of signals. Studies by Gordon Martin confirmed this behaviour when simulated data was used, but showed better performance for MUSIC when actual HF electromagnetic data was used [36].

4.3.7 Other MUSIC variants

Two MUSIC variants intended to permit use with highly correlated multipath arrivals are applicable for uniformly spaced line arrays. One technique uses averaging of the correlation matrices from subarrays [55], [49]. In effect this tends to restore the Toeplitz structure that one would expect with uncorrelated sources. However, it both reduces the effective aperture and the number of arrivals that can be handled by a factor equal to the number of subarrays used. The second technique is called Root MUSIC. For a uniform line array, the steering vectors are sinusoids, so finding steering vectors orthogonal to the noise subspace is equivalent to finding zeros of

polynomials on the unit circle. In Root MUSIC, zeros are rejected if they lie too far from the unit circle, and zeros close to the unit circle are projected onto the unit circle [50]. Root MUSIC is reported to provide better threshold performance than classical MUSIC.

4.3.8 Pisarenko Harmonic Retrieval

Two spectral estimation techniques are both sometimes referred to as "Pisarenko's Method". One is a family of methods based on the eigenvalues and eigenvectors of the correlation matrix. Particular members of this family are chosen by picking a monotone function and its inverse. In this report, we discuss only the Pisarenko special harmonic retrieval method [48]. This method determines the frequencies and amplitudes of sinusoids in additive white noise. Although it preceded (and inspired) MUSIC, it is best understood as a special case of MUSIC, with the following additional assumptions:

1. The noise is white.
2. The data correlation matrix is Toeplitz. In df application this corresponds to uncorrelated sources and a uniformly spaced line array.
3. The noise subspace is treated as having dimension equal to 1.

The Pisarenko method requires finding the eigenvector of a Toeplitz correlation matrix corresponding to the smallest eigenvalue, and then finding sinusoid vectors orthogonal to that eigenvector. Although it can provide high resolution it is not recommended for the following reasons:

1. The Toeplitz assumption makes it inapplicable when even partially coherent multipaths are present.
2. Simulation studies and theoretical analysis have shown that the frequency (or angle in the df application) estimates have large variance or poor statistical efficiency.
3. Gordon Martin's studies using HF data have confirmed the large variance of the direction estimates [36].

4.3.9 ESPRIT

ESPRIT requires an array consisting of two identical subarrays, with the same displacement vector between each element of the first array and the corresponding element of the second array. Alternatively, the total set of array elements may be regarded as a set of M dipoles, with the same

differential displacement vector (length an orientation) for each dipole. The following advantages are claimed for ESPRIT in comparison with MUSIC [45], [46], [47], [50]:

1. ESPRIT eliminates the computationally expensive search through the DOA spectrum.
2. It provides better performance in the presence of highly correlated multipaths.
3. It provides a lower noise threshold.
4. It does not require knowledge of the array manifold.

However, the following disadvantages of ESPRIT for spatial signal processing should also be noted [63]:

1. It is applicable only to a very limited set of array geometries.
2. Even when a two-dimensional array is used, it only produces the same information as other methods using a line array - i.e. arrival angle with respect to the differential displacement axis.
3. If the length of the differential displacement vector exceeds half a wavelength, then ESPRIT will have grating lobe ambiguities.

For applications other than spectrum analysis and direction finding with a uniformly spaced line array spaced closer than spatial Nyquist, the above problems limit the usefulness of ESPRIT.

4.3.10 ESPRIT Variants

The original formulation of ESPRIT was based on solving a non-symmetric generalized eigensystem. Two recent alternatives have been proposed: Total Least Squares ESPRIT, and Orthogonal Procrustes ESPRIT [78],[79],[80], [51]. Total Least Squares ESPRIT is claimed to provide slightly better performance at low signal to noise ratio.

4.3.11 Cadzow's Eigenvector fitting method

Cadzow's method is a modification of MUSIC that avoids the requirement that the source correlation matrix be strictly positive definite [16]. That is, perfectly coherent sources are permitted. The method starts with the eigensystem and "DOA spectrum" calculations as in MUSIC. The noise is assumed to be spatially white, or prewhitened if necessary. Then it is

observed that each eigenvector of the data correlation matrix is a linear combination of steering vectors from the source directions. Source directions are then found by fitting each eigenvector by linear combinations of steering vectors. Usually only a few steering vectors (corresponding to highly correlated arrivals) are used to fit each eigenvector. Since the steering vectors depend nonlinearly on the arrival directions, this is a nonlinear least squares problem, and must be solved iteratively.

4.3.12 Maximum Likelihood Parameter Estimation

This technique (not to be confused with Capon's so-called "maximum likelihood" method) is based on the statistical maximum likelihood parameter estimation method, and assumes that the noise distribution is Gaussian. Maximizing the likelihood ratio for jointly estimating the signal amplitudes and arrival directions is then equivalent to a nonlinear least squares fit of the data vectors by linear combinations of steering vectors - a computational task similar to that required by Cadzow's method. It is known from the statistical literature that maximum likelihood parameter estimation has excellent large sample properties - asymptotically unbiased, consistent, and asymptotically efficient. Recent empirical studies of maximum likelihood direction finding suggest that the small sample behavior is very competitive with other high resolution techniques, and that the method is robust in the presence of highly coherent sources [14], [72], [77]. It also extends in a straightforward way to broadband processing.

4.4 Introduction to Nonstationary Signal Analysis using the Wigner-Ville Distribution

An increasing set of signal processing tasks require an extension of spectral analysis to treat nonstationary functions of a time or space variable. Traditional techniques for the analysis of signals and stationary random processes are based on the fact that the sinusoids are eigenfunctions of every shift-invariant linear operator. This includes both the impulse response of a time-invariant linear filter and the covariance function of a wide-sense stationary random process. The Karhunen-Loeve (K-L) expansion represents an arbitrary random process in terms of the eigenfunctions of its covariance function, and therefore permits an expansion in terms of uncorrelated random variables. However, the K-L expansion does not posess convenient properties for the analysis of signals that have been filtered or modulated.

Recent signal analysis techniques using time-frequency distribution (TFD) functions provide a bridge between the the traditional stationary signal analysis methods and the yet to be developed methods for treating general nonstationary problems. The Cohen [19] class of time-frequency distributions provide a time-varying spectrum with many useful properties for a wide class of signals. Different TFDs in the Cohen class are selected via the choice of a weight function. The choice of a weight function determines which useful properties will be posessed by the resulting TFD, and no one choice provides all of the possible desirable properties.

4.4.1 Principal Notation for Time-Frequency Analysis

t	time
f	frequency
i	imaginary unit (also used as an index)
$s(t)$	real signal
$\mathcal{H}$	Hilbert transform
$\hat{s}(t)$	alternate notation for the Hilbert transform of $s(t)$
A^H	Hermitian transpose of a matrix A
$z(t)$	complex analytic signal
$W_z(t,f)$	Wigner-Ville distribution of $z(t)$
$S_z(t,f)$	short-time Fourier transform of $z(t)$
$K_z(t,\tau)$	bilinear kernel corresponding to $z(t)$
$\star_f$	convolution with respect to frequency
$\star_t$	convolution with respect to time
$\star_{tf}$	two-dimensional convolution in t, f
$m(t)$	window or modulation function
f_0	initial frequency
f_c	carrier frequency
f_s	sampling frequency
α	chirp parameter
T	length of analysis window
$\epsilon(n)$	white noise sequence
$c(i,j)$	estimated covariance matrix
$a(i)$	autoregressive model parameters
C^+	Moore-Penrose pseudoinverse of matrix C
λ	eigenvalue
$\mathbf{v}$	eigenvector
σ	standard deviation (also used for singular value)

4.4.2 The Wigner-Ville Distribution

One TFD that has a large set of the possible useful properties, and that has received extensive study in recent years for application to signal processing is the Wigner-Ville Distribution (WVD) [73],[37]. The WVD of a real signal $s(t)$ that has complex analytic signal $z(t)$ is defined as

$$W_z(t,f) = \int z(t+\tau/2)z^*(t-\tau/2)e^{-i2\pi f\tau}\,d\tau$$

where $z(t) = s(t) + i\hat{s}(t)$ and $\hat{s}(t)$ denotes the Hilbert transform of $s(t)$. The WVD can also be expressed as

$$W_z(t,f) = \int K_z(t,\tau)e^{-i2\pi f\tau}\,d\tau$$

where the kernel $K_z(t,\tau)$ is defined as

$$K_z(t,\tau) = z(t+\tau/2)z^*(t-\tau/2).$$

The WVD of a chirp (a function with constant amplitude and quadratic phase) is distributed on a line through the origin of the time-frequency plane, with slope proportional to the chirp rate. The WVD of the convolution of two signals is the convolution in time of their respective WVDs:

$$W_{z*h}(t,f) = \int W_z(t-u,f)W_h(u,f)\,du$$

These two properties are crucial to a number of applications of the WVD:

- Location of Sources in the Fresnel Region
- Coherent Integration of Lines Emitted by a Moving Source
- Identification of Stationarity Intervals for a Nonstationary Random Process
- Radar Imaging

Since a narrowband signal or signal component in the Fresnel region of a line array produces a quadratic phase shift as a function of position along the array, such a source can be localized in range and bearing via a spatial WVD in the temporal Fourier transform domain. [13].

Estimation techniques using signal subspace concepts can be applied to time-frequency distributions by replacing the Fourier transform of the bilinear kernel by a parametric modelling with respect to the variable τ at each time t [8], [12]. Such a parametric high-resolution modified WVD can be used to enhance the resolution attainable in the spatial localization of a source in the Fresnel region of a line array [74], [66].

4.5 The Windowed Wigner-Ville Distribution

Let s(t) be a real continuous-time signal. The analytic signal associated with s(t) may be expressed as

$$z(t) = s(t) + i\mathcal{H}[s(t)] \tag{4.1}$$

Where $\mathcal{H}$ denotes the Hilbert Transform.

The WVD of a signal s(t) is defined as follows:

$$W_z(t,f) = \int_{-\infty}^{+\infty} z(t+\tau/2)z^*(t-\tau/2)e^{-j2\pi f\tau} \tag{4.2}$$

where z(t) represents the analytic signal defined in (1). The importance of the use of this analytic signal is described in [11].

The properties of the WVD are well described in several comprehensive papers, such as [37] and [10]. In particular, the WVD exhibits a range of peaks along the curve describing the instantaneous frequency law of the signal, which allows frequency tracking of time-varying signals.

Inspection of Equation 4.2 indicates that the evaluation of the WVD is a non-causal operation. Thus, the value of the signal should be known for all time before either the Hilbert transform or the WVD can be evaluated. In many practical cases, the signals considered have either a finite duration or band limited spectrum and the problem is overcome by applying the WVD to a windowed version of the signal z(t) at time t_0 such that $z_m(t,t_0) = z(t)m(t-t_0)$, where $m(t)$ is a time-limited window function with duration T. The WVD of the windowed signal is given by a one-dimensional convolution in the frequency variable, f:

$$W_{z_m}(t,f) = W_z(t,f) \star_f W_m(t,f) \tag{4.3}$$

where $\star_f$ denotes frequency convolution.

Thus the effect of windowing is to smear the WVD representation in the frequency direction only; hence the frequency resolution can be increased by using a longer window without adversely affecting the time resolution. This is a significant improvement over the magnitude squared short-time Fourier transform (STFT) $S_{z_m}(t,f)$ where time resolution is degraded by the window as shown by the following relationship

$$|S_{z_m}(t,f)|^2 = W_z(t,f) \star_{tf} W_m(t,f) \tag{4.4}$$

where $\star_{tf}$ denotes convolution in time and frequency.

For notational convenience, we define K_z and k_m such that:

$$K_z(t,\tau) = z(t+\tau/2)z^*(t-\tau/2) \text{ and } k_m(\tau) = m(\tau/2)m^*(-\tau/2) \quad (4.5)$$

The WVD at any point in time is evaluated by shifting the signal z(t) so that the window is centered about $t = 0$. Therefore, the WVD computation reduces to the evaluation of Equation 4.6:

$$\hat{W}_{z_m}(t,f) = \int_{-T/2}^{+T/2} K_z(t,\tau)k_m(\tau)e^{-i2\pi f\tau}d\tau \quad (4.6)$$

where $\hat{W}_z$ is an estimator of W_z.

Equation (6) shows that the estimation of the WVD reduces to the computation of the FT of a signal kernel $K_z(t,\tau)$ as a function of τ with an analysis window $h(t) = k_m(\tau)$, which is a classical problem of signal analysis. An efficient implementation of (6) which uses fundamental symmetry properties of the WVD is described in [9].

4.6 The Autoregressive WVD (AWVD)

4.6.1 Description of the method

Equation (6) shows that WVD analysis may be thought of as a two-step process:

1. A sequence of kernels $K_z(t,\tau)$ is formed.
2. A Fourier Transform of $K_z(t,\tau)$ is taken, with respect to the variable τ, using the analysis window k_m to reduce spectral leakage.

It is easily shown that for the linear FM signal, $K_z(t,\tau)$ becomes a single complex sinusoid with frequency varying with the chirp parameter α determining the slope [7]. That is, if

$$z(t) = e^{i2\pi(f_c t + \alpha t^2/2)}, \quad (4.7)$$

then the WVD kernel becomes:

$$K_z(t,\tau) = e^{i2\pi(f_c+\alpha t)\tau} \quad (4.8)$$

Since a harmonic process can be generated by an autoregressive (AR) model, the real part of the kernel $K_z(t,\tau)$ at time t may be fitted by the output of an appropriate AR model; the AR coefficients will estimate the

cross-section of the Wigner-Ville Distribution at time t. The real part is taken because the kernel has an Hermitian symmetry and therefore, no additional information is contained in the imaginary part. The analytic signal must however be used in the estimation of the kernel. The expression for the kernel is then:

$$K_z(t,\tau) = s(t+\tau/2)s^*(t-\tau/2) + \mathcal{H}[s(t+\tau/2)]\mathcal{H}[s^*(t-\tau/2)] \quad (4.9)$$

where $\mathcal{H}[.]$ represents the Hilbert Transform. A variety of methods for super-resolution spectral estimation are available for use with this method, based on parametric modelling.

i) Autoregressive modelling

A method of spectral estimation which has been widely accepted in seismic [15] and speech processing [33] is the method of auto-regressive modelling. Makhoul [33] shows that the autoregressive modelling is equivalent to solving the normal equations associated with the linear prediction problem for a finite set of observed samples. For the optimum predictor to use only a finite number of samples of the past, the prediction error filter will have only zeros, and the spectrum being estimated must have only poles. It is the use of this structural information which permits the use of maximum entropy estimators when the process model is appropriate [33].

The traditional approach for implementing this method consists of

1. calculating the sample autocorrelation or autocovariance function
2. solving a linear least-squares problem to find the optimum predictor tap weights
3. Fourier transforming the prediction error filter weights.

This method assumes that the sampled kernel data under analysis can be expressed as a p^{th} order AR model. Then the N point truncated kernel must satisfy the difference equation

$$k_z(n) = \epsilon(n) - \sum_{i=1}^{p} a(i)k_z(n-i) \quad (4.10)$$

where N is odd and

$$\begin{aligned} k_z(n) &= z(t+L)z^*(t-L) \\ L &= n - \left(\frac{N-1}{2}\right), \end{aligned} \quad (4.11)$$

$\epsilon(k)$ is a sample of the noise which is assumed to come from a white noise process; $a(i)$, $i = 1 \ldots p$ are the coefficients of the AR model. (Here, we have assumed a change of variable $n = \tau/2$ before discretizing, as reported in [11].)

Then the covariance AR parameter estimates, $\hat{a}$ and $\hat{\sigma}$, are given by the solution of

$$C\hat{\mathbf{a}} = \hat{\sigma}^2 \mathbf{e}_1 \tag{4.12}$$

where C is the covariance matrix formed by the elements

$$c(i,j) = \frac{1}{N-p} \sum_{s=p}^{N-1} k_z^*(s-i) k_z(s-j) \quad for\ i,j = 0 \ldots p \tag{4.13}$$

$$\mathbf{e}_1 = [1\ 0 \ldots\ 0\ ]^T$$

and

$$\hat{\mathbf{a}} = [1\ \hat{a}(1) \ldots\ \hat{a}(p)\ ]^T \ ;\ (p+1\ \ elements)$$

Equation 4.12 is conceptually solved as follows :

$$\begin{aligned} \mathbf{b} &= C^{-1}\mathbf{e}_1 \\ \mathbf{b} &= \frac{\hat{\mathbf{a}}}{\hat{\sigma}^2} \end{aligned} \tag{4.14}$$

Since $a(0) = 1$, $\hat{\sigma}^2 = \frac{1}{b(0)}$ and $\hat{\mathbf{a}} = \hat{\sigma}^2 \mathbf{b}$.

Unfortunately C^{-1} does not exist if C is singular. For this reason, the Moore-Penrose pseudoinverse, C^+, is used. This may be computed from the nonzero eigenvalues, λ_k, and corresponding eigenvectors, $\mathbf{v}_k$, of C as follows:

$$C^+ = \sum_{s=1}^{M} \frac{1}{\lambda_s} \mathbf{v}_s \mathbf{v}_s^H \tag{4.15}$$

where the eigenvalues are ranked in decreasing magnitude $\lambda_1 \geq \lambda_2 \ldots \lambda_p$ and there are m estimated principle components ($m \leq p$). The value m represents the effective rank of the matrix C.

ii) Singular Value Decomposition

As already stated, we chose not to use the standard normal equation solution to the prediction problem because of its known sensitivity to ill conditioning. The calculations shown in this paper were made using the SVD method proposed by Tufts & Kumaresan [68] and subsequently studied by Minami, Kawaka and Minami [39]. The use of the SVD permits the use of a regularized version of the Moore Penrose generalized inverse in which the

regularization can be controlled by setting the small singular values to zero when they are within a threshold, typically a small multiple of machine precision - macheps (smallest number represented by a given computer which when added to 1 returns something greater than 1). This threshold is a function of the SNR, and should be adjusted accordingly.

The AWVD is implemented by using the singular value decomposition (SVD) of the kernel data matrix to compute a reduced rank principle eigenvector approximation to the inverse covariance matrix for use in the AR modelling process. Since C can be expressed as the product of the observation matrix, K, and its Hermitian transpose, the eigenvectors and eigenvalues of C can be determined from a singular value decomposition of K, i.e.

$$C = K^H K \tag{4.16}$$

where K is the matrix defined by the elements $k(i,j) = K_z(p+i-j)$, for $i = 0 \ldots\ N-1-p$, $j = 0 \ldots\ p$. Singular value decomposition of K gives

$$K = U \Sigma V^H \tag{4.17}$$

where U and V are unitary and Σ is a diagonal matrix with the elements, $\sigma(k,k)$, (called the singular values of K) ranked in decreasing order such that $\sigma(1,1) \geq \sigma(2,2) \ldots \geq \sigma(p,p)$. Then the eigenvectors of C correspond to the columns of V and the eigenvalues of C are given by the squares of the singular values of K.

iii) The Algorithm: Details of the Implementation

For a given value of t, eq (10) can be expressed as: $-K\mathrm{a} + \mathrm{n} = \mathrm{k}$, with the following notation:

$$\mathrm{a} = \begin{bmatrix} a_p(p) \\ \vdots \\ a_1(p) \end{bmatrix}$$

$$\mathrm{n} = \begin{bmatrix} n_p(p) \\ \vdots \\ n_1(p) \end{bmatrix}$$

$$\mathrm{k} = \begin{bmatrix} k_p(p) \\ \vdots \\ k_1(p) \end{bmatrix}$$

$$K = \begin{bmatrix} k(0) & k(1) & k(2) & \dots & k(p-1) \\ k(1) & k(2) & k(3) & \dots & k(p) \\ k(2) & k(3) & k(4) & \dots & k(p+1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ k(N-p-1) & . & . & . & k(N-2) \end{bmatrix} \tag{4.18}$$

K is seen to be a Hankel matrix where the elements of the matrix are the same along the reverse diagonal and whose first row is $\mathbf{k}_z^T$.

In the absence of noise, the estimate a can be given by:

$$\mathbf{a} = -K^{+}\mathbf{k}. \tag{4.19}$$

where K^{+} is the generalized inverse of the matrix K obtained via the reduced rank SVD, i.e. is obtained using the inverse of singular values above the threshhold and the corresponding eigenvectors.

This solution corresponds to a least-squares fit of the data to a model using a finite number p of non-harmonically related sinusoids in noise, where p is the order of the autoregressive model. The parameter p is estimated by either the numerical rank of the data matrix K as computed via the SVD, or is estimated by information theoretic methods. The reader is asked to compare this result with the observation in [34] that the discrete Fourier Transform of a data sequence is equivalent (up to a normalization factor) to a least-squares fit with a model which is composed of a finite number (N, size of the DFT) of harmonically related sinusoids. It is through the inclusion of the knowledge that we are looking for p non harmonically non-related sinusoids instead of N harmonically related sinusoids that the improved (high-precision and accuracy) performance of this Autoregressive Wigner-Ville estimation results.

Selection of the rank, m, can be made by determining the number of the m largest singular values which represent primarily the signal components and are usually well separated from the p-m-1 small values which represent the noise. In practice, to represent the spectrum of a signal which consists of L sinusoids, we need a rank of at least $2L$. Slight overestimation of the rank usually does not affect significantly the resulting spectral estimate. The smallest singular values corresponding to the noise (assumed to be equal for white noise), determine the value of σ.

The window length should be chosen such that it is large enough to contain the majority of the signal, but bounded by its finite duration. The window used is at most the natural effective window $m(\tau)m(-\tau)$ of the kernel occuring in the WVD proccss [8].

The cross-section of the Wigner-Ville Distribution at time t is then estimated by:

$$W(t,f) = \frac{\hat{\sigma}^2}{\left|1 + \sum_{n=1}^{p} \hat{a}(n) e^{-i2\pi f n}\right|^2} \tag{4.20}$$

where $\hat{\sigma}^2$ and $\hat{a}(n)$ are determined from a kernel, k_z, centered about time t, where $t = n/f_s$.

4.6.2 Implementation of the AWVD

Conventional WVD analysis as defined in (2) is performed using a software package based on the algorithm described in [8]. In the new method, the estimated AWVD is calculated as shown in Figure 1 using a peak-finding algorithm to ensure accurate representation of the extremely sharp peaks. Two important features of the implementation are:

1. AWVD performance and speed is improved by performing the modelling on the real part of the kernel since the imaginary part contains redundant information. However the analytic signal must still be used to form the kernel, so that low frequency artifacts are not created [11].
2. The analytic signal must be interpolated before forming the kernel.

4.6.3 Results and discussion

i) Linear FM

The signals considered in this section are FM signals. The linear FM signal can be expressed as:

$$z(t) = e^{j2\pi(f_c t + \alpha t^2/2)} \tag{4.21}$$

for which the WVD kernel becomes:

$$k_z(t,\tau) = e^{i2\pi(f_c + \alpha t)\tau}. \tag{4.22}$$

It is clear that this kernel is composed of one single frequency at any particular time, and that a spectral analysis of the windowed kernel will be more effective than the spectral analysis of the windowed signal since the windowed signal will always contain more than one spectral component.

The test signal is a 512 sample linear FM signal sampled at 200 Hz and modulated from a lower frequency of 10 Hz to an upper frequency of 90Hz. In this implementation, the analytic signal is calculated in the frequency domain and is interpolated by zero padding before performing the inverse discrete time Fourier Transform [11]. The AWVD results are shown in Figure 2 and can be compared with the WVD result in Figure 3.

ii) Hyberbolic FM

The hyperbolic FM signal is expressed as:

$$z(t) = e^{\frac{i2\pi f_0 log(|1+at|)}{a}} \tag{4.23}$$

where f_0 is the initial frequency of the sweep at $t = 0$ and f_c is the center frequency.

Its instantaneous frequency is:

$$f_i(t) = \frac{f_0}{1 + at} \tag{4.24}$$

for which the WVD kernel becomes:

$$K_z(t, \tau) = e^{\frac{i2\pi f_0}{a} \log\left\{ \frac{|1+a(t+\tau/2)|}{1+a(t-\tau/2)|} \right\}}$$

By dividing both numerator and denominator by $2(1 + at)$, the expression for the log can be expanded as:

$$log\left|1 + \frac{a\tau}{2(1 + at)}\right| - log\left|1 - \frac{a\tau}{2(1 + at)}\right|$$

which can be simplified if we have the condition: $\frac{a\tau}{2(1+at)} \ll 1$.

In practice, the signals under analysis are causal and have a finite duration T. Therefore, the maximum value of the numerator in the above inequality is aT. From the expression for $f_i(t)$ given in Equation 4.24, we can also see that the minimum value of the denominator is 1. The previous inequality is therefore valid when $aT \ll 2$ is satisfied. From the expression for $f_i(t)$ in (24), when t equals T, then $f_i(t)$ equals $f_0 + B$. This yields the relation $aT = \frac{B}{(f_0+B)} \ll 2$. By introducing the signal center frequency f_c, this condition can be expressed as:

$$B \ll f_c \tag{4.25}$$

This condition corresponds to narrow band signals. We thus obtain

$$K_z(t, \tau) \simeq e^{i2\pi\left(\frac{f_0}{1+at}\right)} = e^{i2\pi f_i(t)\tau} \tag{4.26}$$

and the same comments as in the linear FM case apply.

results

The AWVD and WVD were applied to an increasing hyperbolic FM signal and the results are shown in figures 4 and 5. The peaks of the AWVD follow the frequency law more closely than the WVD does and the peak width is typically reduced by a factor of 5 to 10. Since the Covariance AR modelling is not subject to windowing distortion, the multiple sidelobes evident in the rectangularly windowed WVD do not appear in the AWVD. Improvement in the frequency law tracking of the AWVD over the conventional WVD is evident. The autoregressive modelling of the WVD kernel data is better than the AR modelling of the signal data because the bilinear form ensures that within a window, we have only one frequency (see equation at the beginning of this section), whereas with the signal data, we always have several frequencies present in the window.

iii) Multicomponent Signals

Although it is expected the the AWVD would perform badly for this class of signals, because of the cross-terms generated by the bilinear kernel, it is nevertheless interesting to present the following example. A multicomponent signal formed from the sum of two chirp signals is analyzed by the AWVD and the WVD. and the results are shown in fig.6 and fig.7. This example shows the cross-product "ghost" peaks which alternate between positive and negative values in the WVD, but are always positive in the AWVD as well as being reduced. The artifacts found in the WVD analysis of multicomponent signals are reduced by the use of the AWVD because the components of the Kernel which give rise to the artifacts are not well modelled by the Autoregressive model which was chosen to model the components of the kernel corresponding to the "autoterms".

The robustness of the method to parameter selection and to additive noise has been tested [12]. For chirp signals analyzed in noise, the results are summarized as follows:

1. For signals with a large SNR, the result is a high accuracy estimate of the WVD, which is largely independent of the model order and rank.

2. The length of the window appears not to affect the resolution of the WVDs.

3. The rank should be approximately the same as (or greater by 1 than) the model order.

4. The higher the model order and rank the greater the resolution of the peaks, but if the model order becomes too large, spurious peaks may occur.

5. The model breaks down below about 5-10 db SNR. This is a reasonable result since the method for estimating the WVD peaks is similar to Prony's method for estimating sinusoids in noise – it can only be successful if the original signal is dominant, *i.e.* the SNR is large enough.

4.6.4 Applications of the AWVD

This section is intended to provide the reader with a few examples of applications of this method, which provided the reason for developing this AWVD estimator.

1. Seismic: As the echolocation search for petroleum resources becomes more difficult, the seismic industry must resolve geological structures with increasing accuracy. Since the bandwidth of the signals is limited by the propagation (attenuation) properties of the earth, modern spectral estimation techniques are playing an ever more important role in seismic prospecting. The high-accuracy AWVD described here provides one method for achieving better identification of fine strata; and this can provide enhanced interpretation of seismic sections, revealing more detail about the thin layers. The modern methods of seismic exploration are now based upon the use of linear FM signals as source instead of classical explosives. These FM signals are emitted by multiple large trucks with vibrating plates in contact with the surface of the ground. This signal fits the model used for the derivation of the AWVD estimator. Under the assumption of the convolution model (weak reflection), then the Wigner-Ville Distribution of the received signal is the convolution (in time) between the WVD of the emitted signal and the WVD of the sequence of reflectors. If the WVD of the sequence of reflectors was known, then using the well-known invertability of the WVD, the sequence of reflectors can be recovered up to a multiplicative constant. Thus the practical problem of Wigner-Ville analysis of the seismic problem is a classical inverse scattering problem in which it is desired to deconvolve the instrument response function, i.e. the WVD of the transmitted signal from the WVD corresponding to the sequence of reflectors. Fourier deconvolution procedures are not appropriate in this case, and high accuracy Wigner-Ville Analysis using the AWVD may be thought of as being equivalent to a numerically robust solution to the classical inverse scattering problem.

2. Sonar/radar problem: The sonar/radar echolocation problem is a generalization of the seismic problem described above in which the indi-

vidual reflectors may also be in motion so that they both provide a delay and a frequency shift of the signal. This problem is normally described in terms of the Ambiguity-function of the transmitted waveform which may be viewed as the output of a matched filter matched to the signal waveform when that waveform is delayed and frequency-shifted. It is well-known [37] that the WVD is equal to the 2DFT of the ambiguity function, and therefore the whole echolocation system can be described in terms of the WVD [74]. In the case where the scattered signals are modelled as Gaussian random variables, the WVD at the output of the receiver may be written as the 2D convolution of the WVD of the transmitted signal with the scattering function of the reflectors in range and doppler. This, too, is now a 2D scattering problem where the instrumentation function, ie; the WVD of the transmitted signal, approximates a line integral, eg. a projection of the scattering distribution when the transmitted signal is a linear FM, and thus the equivalent deconvolution problem is anhomorphic (there is different accuracy in range and doppler). Also, the deconvolution is enhanced by high-accuracy techniques which make the line integral approximation more closely valid, so that the deconvolution can be achieved by tomographic reconstruction techniques [74].

3. Fresnel region beamforming: The third example is an example of the passive localisation problem where multiple spherical waves are received by a uniformly spaced line-array of sensors such as geophones, hydrophones or electromagnetic antennas. It has been shown in the array processing literature that when the receiving array is in the far field, the spherical waves may be approximated by plane waves. Then the distribution of phase along the sensors will be linear and different ratios of phase shifts correspond to different angles of arrival; and thus the problem of estimating the direction of arrival is equivalent to the spectral analysis of the time signal corresponding to the spatial samples of the signal. This type of signal processing is equivalent to decomposing the incoming signal field into a frequency-wavenumber spectrum [28]. When the sources are in the Fresnel zone of the array, then the curvature of the spherical wavefront must be properly accounted for. Inspection shows that it is now a linear FM, which again corresponds to the signal for which the proposed AWVD method will be most effective [13][66]. Although WVD analysis allows for the simultaneous estimation of direction of arrival and range, a high-accuracy WVD is required to provide practical estimation of the range.

LIST OF FIGURES AND CAPTIONS

4.7 Radar Imaging

A new formulation of tomographic imaging for rotating targets is presented in terms of the Wigner-Ville distribution. This new formulation both simplifies the signal processing required for image generation and is applicable in both mono-static and bi-static situations. Also, the new formulation will allow the use of parametric modifications of the WVD, including the AWVD described above, to improve the resolution attainable in radar imaging.

The target is assumed to have multiple scatterer within each range-Doppler resolution cell so that the received signal will be a sample function of a Gaussian random variable. When the transmitted waveform is a linear frequency modulated chirp then the expected value of the output of a matched filter receiver is the two-dimensional convolution of the range-Doppler scattering function of the target with the auto-ambiguity surface of the transmitted waveform. Under the additional assumption that the transmitted signal has a Gaussian or rectangular envelope then the receiver can be simplified to a short-time Fourier transform receiver using only *a priori* information about the envelope and the chirp rate of the transmitter. This simplification eliminates the need for detailed phase information at the receiver and thus bi-static operation is feasible. Since most surfaces are rough at optical frequencies this form of signal processing should be particullarly useful with optical sensors such as coherent laser radars (LIDARs).

4.7.1 Introduction to Range-Doppler Imaging

Range-Doppler imaging (scatterer distribution mapping) is a general class of inverse problem which arises in remote sensing applications in either radar or sonar whenever the sensor is in relative motion with respect to the object or objects under study. In principle the sensors can radiate either electromagnetic (radar) or acoustic (sonar) signals. The radiated signals are usually in the form of a modulation on a carrier wave. The ratio of modulation bandwidth to carrier frequency is called the fractional bandwidth. We can distinguish between two cases, depending on the fractional bandwidth:

1. Small Fractional Bandwidth: The fractional bandwidth is usually small in the case of radar and thus motion induced Doppler can be treated as a single frequency shift of the entire modulation spectrum.

2. Large Fractional Bandwidth: The fractional bandwidth is often large in the case of sonar and thus motion induced Doppler should be treated as a change in time scale of the modulation signal or equivalently the Doppler shift caused by the motion is proportional to the frequency of the components of the transmitted spectrum.

For simplicity we will consider only the small fractional bandwidth case and will use radar terminology and examples. However, the signal processing concepts presented in this section can be extended to the large fractional bandwidth case for use in sonar signal processing.

The problem which will be the subject of this paper is the mapping of rigid objects which are in motion with respect to a stationary sensor. Although rigid objects are the primary targets of interest in most remote sensing applications non rigid scatter distributions such as a flight of birds or a school of fish may be mapped if the overall shape of the swarm does not change too much during the mapping operation. By specalizing to rigid objects, it is possible in many applications to incorporate *a priori* or in some cases *a posteriori* information about the target and thus convert the measured scatterer distribution to a range and cross-range image of the object. Two types of scattering, deterministic and stochastic, have been considered in the literature. This paper will assume the stochastic model of scattering which is usually valid when the target is rough with respect to the wavelength of the illuminations. Thus the inverse problem is to estimate the range-Doppler scattering function of the target which represents the average power reflected from a resolution cell of the target. If the target is essentially two-dimensional, (i.e. a rotating disk), when viewed along the line of sight (LOS) of the radar than the range-Doppler

cells of the scattering function form a one-to-one mapping with range and cross-range cells of the object. If the target is essentially three-dimensional, (i.e. a rotating sphere), then the mapping is no longer one-to-one and in general there may be two or more range and cross-range cells with the same range and Doppler (e.g. the rotation axis of the sphere is perpendicular to the LOS of the radar).

Although the analysis of the tomographic imaging method in this paper is stated in terms of a monostatic radar, (i.e. transmitter and receiver either co-located or closely located), the analysis can be extended to a bi-static radar, (transmitter and receiver at different locations). In the bi-static case the transformation from range and Doppler to range and cross-range is more difficult due to the geometrical terms in the transformation [3]. However, the measurement of the range-Doppler scattering function is still simple when the analysis is performed using the Wigner-Ville distribution instead of the ambiguity surface, (i.e. magnitude squared of the ambiguity function). As will be subsequently shown, the estimation of the scattering function separates into a measurement of the Wigner-Ville distribution of the received signal made at the receiver and the utilization of a prori information about the transmitted signal.

Although the concept presented in this paper is based on a stochastic model of the reflection process, it still works as described even when the target is deterministic and not fluctuating. In this latter case the signal processing proposed in this paper may be simplified by reducing the number of different chirp signals used and thus may be useful at microwave frequencies where longer coherent integration times are required due to the reduced Doppler sensitivy of the lower carrier frequency, compared to the LIDAR case. Also, since the nominal resolution of the range-Doppler cell is proportional to the bandwidth (range resolution) and the duration of coherent integration (Doppler resolution) use at microwave frequencies may require stepped frequency chirp signals instead of continuous chirp signals to satisfy the bandwidth limitations of available components [71].

4.7.2 WVD Background for Echography

In order to understand the role of the Wigner-Ville distribution in radar/sonar signal processing, it is necessary to introduce a model of the received signal which is sufficiently realistic to be practical yet sufficiently simple so that it can be easily understood. Therefore, only scalar signals will be considered. This assumption is true for acoustic waves propagating in a liquid medium and is a simplifying assumption for electromagnetic waves.

The received signal will be modeled as the linear superposition of scaled versions of the transmitted signal delayed by the round trip time and mod-

ified by Doppler. Since the transmitted signal has only limited range and Doppler resolution then the received signal may be simplified by assuming that the scale factor is the product of a constant that depends on the range and a complex random variable which acounts for interference between unresolved target scatterers and propagation multipaths. A reasonable assumption, when the individual contributions are sufficiently numerous and approximately equal in magnitude with uniformly distributed phase, is that the random variable can be modeled as a zero mean Gaussian with variance $\sigma(t, f)$ which depends on the range and Doppler. This function is traditionally called the scattering function and this target model has been described by VanTrees [69].

The scattering function will be assumed to be uncorrelated in frequency due to orthogonality of the Doppler shifts for narrow band signals. However, different delays may be correlated depending on the characteristics of the target. In order to be able to include both deterministic as well as stochastic variation from pulse to pulse the random variable will be modeled with an exponential time correlation which can be generated by a recursive filter. Conventional processing only works well when the correlation is high, (i.e. the target scattering is essentially deterministic). Speckle noise is associated with single look SAR images.

4.7.3 Comparision of Radar Imaging Techniques

Four different analyses of the imaging of radar targets have been proposed in the literature.

In the first and simplest the received signal from a spatially distributed target is illuminated at successive frequencies as a function of angle. If the scattering is assumed to be deterministic then the *two-dimensional Fourier transform* of the received signal gives the range and cross-range components of the scattering distribution provided the target is in the far field of the illumination and the variation in angle is sufficiently small that a rectangular approximation of a polar grid is valid. This type of signal processing is often used in radar ranges where the angular orientation can be measured and the signal to noise ratio is sufficiently large that matched filter processing is not required.

In the second method *matched filtering* is used with either continuous or stepped frequency chirp signals. Succesive range compressed signals from a moving target are aligned in time in order to compensate for the variation in range between succesive measurements and then a discrete Fourier transform is computed on the complex matched filter outputs for successive measurements. This Fourier transform computes the Doppler frequency as a function of range and is related to the cross-range of a rigid object if the

effective azimuth rotation rate relative to the line of sight (LOS) is known. This method is effective for deterministic signal returns, such as those from metallic aircraft illuminated by microwave frequencies. However, when the target is rough relative to the illuminating wavelength then the scattering is stochastic and this type of processing is not appropriate.

In the third method the *cross-ambiguity function* between the transmitted and the received signal is computed. For deterministic scattering distributions this integral is proportional to the two-dimensional Fourier transform of the product of a complex reflection coeficient with the ambiguity function of the transmitted waveform–but where the proportionality factor is complex and range dependant. This method of range-Doppler imaging has been analyzed by Feig and Grunbaum and leads to a new form of tomographic imaging [22].

In the fourth method, the *magnitude squared of the cross-ambiguity function* between the transmitted and the received signal is computed. For stochastic scatterers the expected value of the magnitude squared of the cross-ambiguity function is proportional to the two-dimensional convolution of the range- Doppler scattering function with the magnitude squared of the auto-ambiguity function of the transmitted signal. Thus for rotation relative to the LOS the effective Doppler of the target becomes coupled to the range of the target through the ambiguity function's magnitude squared response. The latter is sometimes referred to as the ambiguity surface. For narrowband signals of unit energy the ambiguity surface is everywhere positive and has unit volume. Thus it may be considered to be a probability density function which distributes the energy from the scatterer in both range and Doppler acording to the particular wave form used by the transmitter. This method of imaging is analagous to the medical method of time-of-flight positron emmision tomography (TOFPET) when finite duration chirp signals are used and is analogous to the medical method of computerized axial tomography (CAT) in the limit of infinite duration chirp signals. The analysis of this method of signal processing will be considered next.

4.7.4 Tomographic Radar Imaging using Ambiguity Surfaces

In 1984, Bernfeld proposed the use of a computerized axial tomography (CAT) analogy to process linear fm synthetic aperture (SAR) images [5]. Instead of transmitting repeatedly the same linear fm signal as the radar moves along its path, Bernfield suggested that chirps of different slopes be transmitted. For a stochastic Gaussian assumption of the scattering process the output of a matched filter receiver may be modeled as the two-

dimensional convolution of the range-Doppler scattering function with the auto-ambiguity surface of the transmitted signal. In the limit of an infinite duration linear fm signal and no additive noise the output of the receiver may be considered to be equivalent to a sequence of line integrals through the scattering function where the orientation of the particular line integral relative to the scattering distribution is determined by the chirp rate of the transmitted signal. The ambiguity surface of an infinite duration linear fm signal may be considered to be a delta function with an orientation in range and Doppler determined by the chirp rate of the signal. Therefore, the convolution of this ambiguity function with the scattering distribution may be considered to be a sequence of tomographic projections of the scattering distribution.

In most sonar and radar applications, infinite duration signals are not practical.Therefore, the tomographic signal processing technique proposed by Bernfield should be modified to accomodate finite duration signals. For a finite duration linear fm signal the ambiguity surface is only an approximation to a line integral. In 1985 Snyder and Whitehouse proposed a signal processing modification based on the medical method of time-of-flight positron emission tomography (TOFPET) [58].

When a linear fm signal with a Gaussian envelope is used as the transmitted signal in an echo ranging system then the auto-ambiguity surface of the transmitted signal is an asymmetric two-dimensional Gaussian function. The inverse problem which results when a number of different orientations of the two- dimensional Gaussian are convolved with an unknown but bounded two-dimensional distribution has been studied in the context of reconstructing medical radio nucleide activity distributions [57]. One of the techniques of solving this inverse problem is called the "confidence weighting" algorithm and is a generalization of the CAT technique of back projection and superposition followed by two-dimensional filtering.

A phenomological description of confidence weighting can be easily constructed to motivate the confidence weighting tomographic algorithm. In the conventional CAT algorithm there is no information in the projection or line integral about the location of the contributions to the line integral. Therefore, the value of the line integral is uniformly distributed along the path of the line integral, (i.e. back projection). In the case of TOFPET or the proposed tomographic radar imaging algorithm the pointspread function of the tomograph or the ambiguity surface provides *a priori* information about the contributions to the receiver output. Since the normalized ambiguity surface has unit volume and is positive, it may be interpreted as a probability distribution and thus confidence weighting is simply a method analogous to back projection for incorporating the *a priori* knowledge.

4.7.5 WVD Interpretation of Imaging

It is well known that the Wigner-Ville distribution (WVD) and the ambiguity function of Sussman are related by a two-dimensional Fourier transform [67]. Thus it is not surprising that radar or sonar detection and estimation can be reinterpreted in terms of the WVD. However, it is not without some difficulty as will be presented in the following section. But the difficulties are minor compared to the insight provided by reformulating the signal processing equations. In particular, such an interpretation of the signal processing for bi-static operation of the radar or sonar is easier to understand, since the effect of receiver mismatch may be predicted and in the case of linear frequency modulated (LFM) signals simplified signal processing either in terms of the short time Fourier transform (STFT) or STRETCH processing is possible.

The WVD has many interesting properties [37]. However, only a few will be needed for our purposes. These are:

- Realness
- The shift properties
- Moyal's formula
- The marginal properties
- Relationship to the ambiguity function
- Relationship to the ambiguity surface

Unfortunately, there are some additional properties which would be desirable in a time-frequency distribution, but which are attainable only when other desirable proretties are sacraficed. These are:

- Linearity; Must give up realness or resolution.
- Positivity; Must give up resolution,the marginal properties, and the general applicability of Moyal's formula.

Tomographic signal reconstruction techniques restores linearity and as well as positivity. High resolution Wigner-Ville analysis shows promise in restoring the resolution [12]. However, Moyal's formula is no longer applicable, and therefore there is a small loss of detectability which is not usually important for imaging applications [23].

The tomographic processing algorithm using the Wigner-Ville distribution can be simply deduced from the corresponding tomographic algorithm

using the ambiguity surface. The expected value of the cross-ambiguity surface of the matched filter receiver is the two-dimensional convolution of the scattering function with the auto-ambiguity function of the transmitted waveform [69]. Using the relationship between the ambiguity surface and the Wigner-Ville distribution the cross-ambiguity surface is rewritten as the two-dimensional convolution of the Wigner-Ville distribution of the received signal with the Wigner-Ville distribution of the transmitted signal. In a simillar manner the auto-ambiguity surface of the transmitted signal is the two-dimensional convolution of the Wigner-Ville distribution with itself. If the two-dimensional Fourier transform (i.e. ambiguity function) of the WVD of the transmitted signal is positive then the an equivalent formulation is that the expected value of the Wigner-Ville distribution of the received signal is the two-dimensional convolution of the target scattering function with the Wigner- Ville distribution of the transmitted signal.

LASER detection and ranging (LIDAR) technology has significant resolution advantages over conventional microwave radar techniques. The technology has received considerable attention as a means of imaging spaceborne objects using inverse synthetic aperture radar (ISAR) processing techniques using range resolved tomographic reconstruction. These ISAR techniques, however, become ineffective when the object is only viewed for a small fraction of a rotation, a common occurrence in strategic defense applications. Furthermore, the formidable signal processing requirements inherent in a ISAR imaging application require innovative techniques to minimize the size, weight, and power of the signal processor, particularly for spaceborne systems. Recently, Whitehouse and Boashash have suggested that a LIDAR imaging processor which uses tomographic reconstruction techniques to form an image may overcome the limited viewing angle problem associated with range resolved tomographic ISAR processing [75]. In this new technique, the chirp rate of the LIDAR transmitted pulse is varied during the illumination period. For these signals, the Wigner-Ville distribution, which simultaneously measures the signal time-delay and Doppler shift, is known to be an approximation to a knife-edge whose orientation in the time-delay, Doppler plane is a function of the chirp rate. Thus the output of a Wigner-Ville receiver which is the convolution of this rotated function with the received LIDAR signal and corresponds to calculating a projection of the object's scattering distribution. Time of flight positron emmision tomography (TOFPET) reconstruction techniques can then be employed to reconstruct an image of the object by confidence weighting, (i.e. convolving the received Wigner-Ville distribution with the Wigner-Ville distribution of the illuminating signal). and filtering the superposition of the confidence weighted signals.

4.8 Conclusions

The direction finding algorithms we have considered impose a substantial computational load for their real-time implementation. Most of this computational load consists of matrix operations:

- matrix-vector multiplication
- linear equation solution
- orthogonal triangularization
- triangular backsolve
- Hermitian symmetric eigensystem solution
- singular value decomposition
- Hermitian symmetric generalized eigensystem solution

These matrix operations also comprise a substantial part of the computational load for several useful methods for data compression, spectral analysis, and pattern recognition.

4.8.1 Research Issues in Parameter Estimation

Two of most promising parameter estimation techniques for further study for application to direction finding are Cadzow's method and the Maximum Likelihood method. MUSIC and MVDR should be viewed as baseline algorithms for comparison with the newer methods. Both Cadzow's eigenvector fitting method and the Maximum Likelihood Method permit high resolution direction finding in the presence of fully coherent multipath, and both are applicable to very general array geometries. However, both require the solution of a computationally expensive nonlinear least squares problem. At present, it is not only difficult to specify a preferred method for solving this nonlinear least squares problem, but even to specify the number of arithmetic operations needed for an acceptable solution. In order to make Cadzow's method and MLM feasable for real-time implementation and examine their robustness, further study is needed for:

- Convergence of Nonlinear Least Squares Algorithms
- Parallelization of Nonlinear Least Squares Algorithms
- NLS solution accuracy needed for Cadzow's Method and MLM

- Effect of element gain, phase, and position errors on Cadzow's Method and MLM

For the special case of radar imaging via estimation of the target scattering function, time-frequency analysis via the application of signal subspace techniques to the Wigner-Ville distribution leads to problems analogous to spectral estimation [76]. For such problems, the spectral analysis of the WVD's bilinear kernel is analagous to direction finding using a uniformly spaced line array, and ESPRIT and its variants are strong candidate methods. These techniques will require parallel implementations for:

- The Nonsymmetric Generalized Eigenvalue Problem
- Total Least Squares
- The Orthogonal Procrustes Problem

4.8.2 Parallel Computation

Surveys of parallel algorithms and architectures for many of the matrix computations discussed in this paper are available [62], [64]. Although nonlinear least squares algorithms use the preceding operations, the resulting iterative algorithms are not as well understood as the classical matrix operations. Additional research is also needed in the area of parallel algorithms for the nonsymmetric eigenvalue problem, since current techniques using parallel Jacobi-like methods for computing the Schur decomposition can have convergence problems for non-normal matrices. [20]

Acknowledgement

One of the authors (HJW) wishes to thank Prof. Donald Snyder of Washington University for many stimulating discussions of tomographic techniques.

One of the authors (BB) wishes to acknowledge the support of the Australian Research Council.

One of the authors (JMS) wishes to acknowledge the support of the NOSC Internal Research Program.

Bibliography

[1] H. C. Andrews and C. L. Patterson, "Outer Product Expansions and Their Uses in Digital Image Processing", *American Math. Monthly,* Vol.82, No.1, January 1975, pp. 1-13

[2] H. C. Andrews and C. L. Patterson, "Singular Value Decomposition (SVD) Image Coding", *IEEE Trans. on Communications,* April 1976, pp. 425-432

[3] O. Arikan and D.C. Munson, Jr., "A tomographic formulation of bistatic synthetic aperture radar," in W.A Porter and S.C. Kak, Editors, *Proceedings of the 1988 international conference on advances in communication & control systems*, Dept. of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803, October 19-21, 1988, Baton Rouge, LA, pp. 418-431

[4] William P. Ballance, "The Singular Value Decomposition: An Introduction", *Conference Record of the eighteenth Asilomar Conference on Circuits, Systems & Computers,* held at Pacific Grove, California, Nov.5-7, 1984, IEEE Catalog No.85CH22200-4, pp. 148-157

[5] M. Bernfeld, Chirp Doppler radar, *Proc. IEEE Lett.*, Vol 72, No.4, April 1984, pp. 540-541

[6] G. S. Bienvenu, "Influence of the Spatial Coherence of the Background Noise on High Resolution Passive Methods", *ICASSP 79 Record*, pp. 306-309

[7] Boashash, B. and H.J. Whitehouse, "Seismic Applications of the Wigner-Ville Distribution", *Proceedings of IEEE International Symposium on Circuits and Systems*, Vol. 1 of 3, pp. 34-37, San Jose, USA, May 1986

[8] B. Boashash and H.J. Whitehouse, "High Resolution Wigner-Ville Analysis", *Onzième Colloque GRETSI*, Nice, 1-5 June, 1987, pp. 205-208

[9] Boashash, B. and P. Black, "An efficient real time implementation of the Wigner-Ville Distribution", *IEEE transactions on Acoustics, Speech and Signal Processing*, Vol. 35, pp. 1611-1618, November, 1987

[10] Boashash, B. "Time-Frequency Signal Analysis and Synthesis - the choice of a Method and its Application", *Proc. of SPIE Conference, Advanced Algorithms and Architectures*, San Diego, Vol. 975, pp. 164-185, Aug. 1988

[11] Boashash, B., "Note on the Use of the Wigner Distribution for Time Frequency Signal Analysis", *IEEE Trans. on ASSP*, Vol. 36, pp. 1518-1521, September, 1988

[12] Boashash, B. and H. J. Whitehouse, "High Accuracy Time-Frequency Analysis of FM Signals using an Autoregressive Wigner-Ville Distribution", submitted to the *IEEE Transactions on Acoustics, Speech, and Signal Processing*

[13] B. R. Breed and T. E. Posch, "A Range and Azimuth Estimator Based on Forming the Spatial Wigner Distribution", *Proc. ICASSP '84*, paper 41B9

[14] Y. Bressler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, No.5, Oct. 1986, pp. 1081-1089

[15] Burg, J.P., "Maximum Entropy Spectral Analysis", *Proc. 37th Meet. Soc. of Exploration Geophysicists*, Oklahoma City, OK, 1967

[16] J. A. Cadzow, "A High Resolution Direction-of-Arrival Algorithm for Narrow-Band Coherent and Incoherent Sources", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol.36, No.7, July, 1988, pp. 965-979

[17] J. Capon, R. Greenfield, and R. J. Kolker, "Multidimensional Maximum-Likelihood Processing of a Large Aperture Seismic Array", *Proc. IEEE*, Vol.55, p.192- 211, Feb. 1967

[18] J. Capon, "High Resolution Frequency-Wavenumber Spectrum Analysis", *Proc. IEEE*, Vol. 57, pp. 1408-1418, Aug.1969

[19] Cohen, L., "Time-Frequency Distributions: A Review", *Proc. IEEE*, July 1989

[20] P. J. Eberlein, "On the Schur Decomposition of a Matrix for Parallel Computation", *IEEE Trans. on Computers*, Vol. C-36, No.2, Feb. 1987, pp. 167-174

[21] M. Feder and E. Weinstein, "Optimal Multiple Source Location Estimation via the EM Algorithm", *ICASSP 85 Proceedings*, Vol.4, paper 45.6, pp. 1762-1765

[22] E. Feig and F.A Grünbaum, "Tomographic methods in range-Doppler radar," *Inverse Problems*, Vol. 2, pp. 185-195, 1986

[23] P. Flandrin, "A Time-Frequency Formulation of Optimum Detection", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol.36, No.9, September 1988, pp. 1377-1384

[24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1983, third printing, with corrections, 1985

[25] I. J. Good, "Some Applications of the Singular Decomposition of a Matrix", *Technometrics*, Vol.11, No.4, November 1969, pp. 823-831

[26] S. Haykin (Ed.), *Nonlinear Methods of Spectral Analysis*, Springer-Verlag, 1979

[27] D. H. Johnson and S. R. De Graaf, "Improving the Resolution of Bearing in Passive Sonar Arrays by Eigenvalue Analysis", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-30, No.4, pp. 638-647, August 1982

[28] Kailath, T.(Ed.), *Modern Signal Processing*, Hemisphere Publishing Corporation, 1985

[29] F. T. Luk, "A Parallel Method for Computing the Generalized Singular Value Decomposition", *Journal of Parallel and Distributed Computing*, Vol2, 1985, pp. 250-260

[30] M. T. Ma, *Theory and Applications of Antenna Arrays*, John Wiley & Sons, 1974

[31] R. N. McDonough, "Application of the Maximum-Likelihood Method and Maximum Entropy Method to Array Processing", Chapter 6 of *Nonlinear Methods of Spectral Analysis*, edited by S. Haykin, Springer-Verlag, 1979

[32] J. G. McWhirter and T. J. Shepherd, "Efficient minimum variance distortionless response processing using a systolic array", *Proc. SPIE*, Vol.975, paper 975-39, pp. 385-392

[33] Makhoul, J., "Linear Prediction: A Tutorial Review", *Proc. IEEE*, Vol. 63, No.4, pp. 561-580, April, 1975

[34] Marple, S.L., *Digital Spectral Analysis - with applications*, Prentice Hall Inc, 1985

[35] S. L. Marple, *Digital Spectral Analysis*, Prentice-Hall, 1987

[36] G. E. Martin, " Signal subspace processing of experimental radio data", *SPIE Proceedings* Vol.975, paper 975-11, pp. 101-107, 1988

[37] W. Mecklenbräuker, "A Tutorial on Non-Parametric Bilinear Time-Frequency Signal Representations", in *Les Houches, Session XLV, 12 Août-6 Septembre 1985, Traitment du Signal / Signal Processing*, edited by J. L. Lacoume, T. S. Durrani, and R. Stora, Elsevier Science Publishers, B. V., 1987, pp. 277-336

[38] M. I. Miller and D. R. Fuhrmann, "An EM Algorithm for Direction-of-Arrival Estimation for Narrowband Signals in Noise", *Proc. SPIE*, Vol.826, pp. 101-103 (1987)

[39] Minami, K., Kawata, S. and Minami, S., "Superresolution of Fourier Transform Spectra by Autoregressive Model Fitting with Singular Value Decomposition", *Appl. Optics*, Vol. 24, no. 2, pp. 162-167, Jan. 1985

[40] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*, John Wiley & Sons, 1980

[41] J. C. Nash, *Compact Numerical Methods for Computers*, John Wiley & Sons, New York, 1979

[42] U. Nickel, "Angle Estimation with Adaptive Arrays and Relation to Superresolution", *IEE Proc.*, Vol.134, Part II, pp. 77-82 (1987)

[43] N. L. Owsley, "Sonar Array Processing", Chapter 3 of *Array Signal Processing*, edited by S. Haykin, Prentice-Hall, 1985

[44] C. C. Paige, "Computing the Generalized Singular Value Decomposition", *SIAM J. Sci. Stat. Comput.*, Vol.7, No.4, October 1986, pp. 1126-1146

[45] A. Paulraj, R. Roy, and T. Kailath, "Estimation of Signal Parameters via Rotational Invariance Techniques - ESPRIT", *Conference Record of the Nineteenth Asilomar Conference on Circuits, Systems, and Computers*, Nov.6-8, 1985, pp. 83-89

[46] A. Paulraj, R. Roy, and T. Kailath, " A subspace rotation approach to estimating directions of arrival", *Proc. ICASSP 86*, April 7-11, 1986, Vol.4, paper 47.2, pp. 2495-2498

[47] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation", *Proc. IEEE*, July 1986, pp. 1044-1045

[48] V. F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function", *Geophys. J. R. astr. Soc.*, Vol.33, pp. 347-366, (1973)

[49] S. S. Reddi, "On a Spatial Smoothing Technique for Multiple Source Location", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-35, No.5, May, 1987, p.709

[50] R. Roy, A. Paulraj, and T. Kailath, "Comparative Performance of ESPRIT and MUSIC for Direction-of-Arrival Estimation", ICASSP 87, IEEE Order No. CH2396-0/87, Volume 4, paper 54.12, pp. 2344-2347

[51] R. Roy and T. Kailath, "Invariance Techniques and high-resolution null steering", *Advanced Algorithms and Architectures for Signal Processing III*, SPIE Vol. 975, (1988), paper 975-35, pp. 358-367

[52] R. O. Schmidt " Multiple Emitter Location and Signal Parameter Estimation", *Proceedings of the RADC Spectrum Estimation Workshop*, Rome Air Development Center, pp. 243-257, Oct.3-5, 1979

[53] R. O. Schmidt, *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*, Ph.D. Dissertation, Dept. of Electrical Engineering, Stanford University, November, 1981

[54] R. Schreiber and P. J. Kuckes, "Systolic Linear Algebra Machines for Digital Signal Processing", in *VLSI and Modern Signal Processing*, edite by S. Y. Kung, H. J. Whitehouse, and T. Kailath, Prentice Hall, 1985, pp. 389-405

[55] T. J. Shan, M. Wax, and T. Kailath, "On Spatial Smoothing for Direction of Arrival Estimation of Coherent Signals", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol.ASSP-33, No.4, Aug. 1985, pp. 806-811

[56] S. D. Silverstein, "Robust Spectral Estimation: Autocorrelation Based Minimum Free Energy Method", *Twenty-Second Asilomar Conference on Signals, Systems & Computers*, Oct. 31-Nov. 2, 1988

[57] D. L. Snyder, L. J. Thomas, Jr., and M. M. Ter-Possian, "A Mathematical Model for Positron Emission Tomography Systems Having Time-of-Flight Measurements", *IEEE Trans. on Nuclear Science*, Vol. NS-28, pp. 3575-3583, June 1981

[58] D.L. Snyder, H.J. Whitehouse, J.T. Wohschlaeger, R.C. Lewis, "A New Approach to RADAR/SONAR Imaging", *Proceedings of the SPIE 30th Annual Technical Symposium*, San Diego, CA, August 1986

[59] J. M. Speiser, H. J. Whitehouse, and N. J. Berg, " Signal Processing Architectures using Convolutional Technology", *Proc. SPIE* Vol.154, pp. 66-80, 1978

[60] Speiser, J., Bromley, K., and Whitehouse, H.J., "Signal Processing Applications for Real-Time Matrix Operations", in *Advanced Computer Concepts*, LJI Conference Proceedings, La Jolla Institute, Editor J.C. Solinsky, pp. 35-46, 1981

[61] J. M. Speiser, "Progress in Eigenvector Beamforming", *SPIE Proceedings* Vol.564, paper 564-01, 1985

[62] J. M. Speiser, "Signal Processing Computational Needs", *Proc. SPIE*, Vol.696, pp. 2-6 (1986)

[63] J. M. Speiser, "Some Observations Concerning the ESPRIT Direction Finding Method", *Proc. SPIE*, Vol.826, pp. 178-185, (1987)

[64] J. M. Speiser, "An Overview of Matrix-Based Signal Processing", *Conference Record of the Twenty-First Asilomar Conference on Circuits, Systems, and Computers*, held at Pacific Grove, CA, Nov.2-4, 1987, IEEE Computer Society Order No. 760, pp. 284-289

[65] B. D. Steinberg, *Principles of Aperture and Array System Design*, John Wiley & Sons, 1976

[66] A. L. Swindlehurst and T. Kailath, "Near-field source parameter estimation using a spatial Wigner distribution approach", *Advanced Algorithms and Architectures for Signal Processing III*, SPIE Vol. 975, (1988), paper 975-09, pp. 86-92

[67] H.H Szu, "Two-dimensional optical processing of one-dimensional acoustic data", *Optical Engineering* Vol. 21, 1982, pp. 804-813

[68] Tufts, D.W. and Kumaresan, R., "Singular value decomposition and improved estimation using linear prediction", *IEEE Trans ASSP*, vol.30, no.4, Aug 1982

[69] H.L. Van Trees, ***Detection, Estimation, and Modulation Theory***, Part III, Radar-Sonar Signal Processing and Gaussian Signals in Noise, John Wiley and Sons, New York, 1971

[70] R. A. Wagstaff and A. Baggeroer (editors), *High Resolution Spatial Processing in Underwater Acoustics*, Naval Ocean Research and Development Activity, NSTL

[71] D.R. Wehner, *High Resolution Radar*, Artech House, 1987

[72] A. J. Weiss, A. S. Willsky, and B. C. Levy, "Maximum Likelihood Array Processing for the Estimation of Superimposed Signals", em Conference Record of the Twenty-First Asilomar Conference on Signals, Systems, and Computers, held at Pacific Grove, CA, Nov.2-4, 1987, Vol.2, pp. 845-848

[73] H. J. Whitehouse and B. Boashash, "Signal Processing Applications of Wigner-Ville Analysis", *SPIE Vol.696 Advanced Algorithms and Architectures for Signal Processing (1986)*, pp. 156-162

[74] Whitehouse, H. J., "Interpretation of radar imaging in terms of the Wigner-Ville distribution", presented at *Advanced Algorithms and Architectures in Signal Processing*, SPIE International Symposium, San Diego, Calif, August 15-17, 1988

[75] H. J. Whitehouse and B. Boashash, "Delay-Doppler Radar/Sonar Imaging", presented at EUSIPCO 88, Grenoble, France, 5-9 Sept, 1988

[76] H. J. Whitehouse and L. B. White, "Resolution Enchancement in Time-Frequency Signal Processing Using Tomographic Imaging", to appear in *ICASSP 89*, May 23-26, 1989, Glascow, Scotland

[77] I. Ziskind and M. Wax, "Maximum Likelihood Localization of Multiple Sources by Alternating Projection", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-36, No.10, Oct. 1988, pp. 1553-1560

[78] M. D. Zoltowski, "Novel Techniques for Estimation of Array Signal Parameters Based on Matrix Pencils, Subspace Rotations, and Total Least Squares", *ICASSP 88 Conference Proceedings*, Vol.V, pp. 2861-2864

[79] M. D. Zoltowski, "Generalized Minimum Norm and Total Least Squares with Applications to Array Signal Processing", *Advanced Algorithms and Architectures for Signal Processing III*, SPIE Vol. 975, (1988), paper 975-08, pp. 78-85

[80] M. D. Zoltowski and D. Stavrinides, "Sensor Array Signal Processing Via a Procrustes Rotations Based Eigenanalysis of the ESPRIT Data Pencil", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 37, No.6, June 1989, pp. 832-861

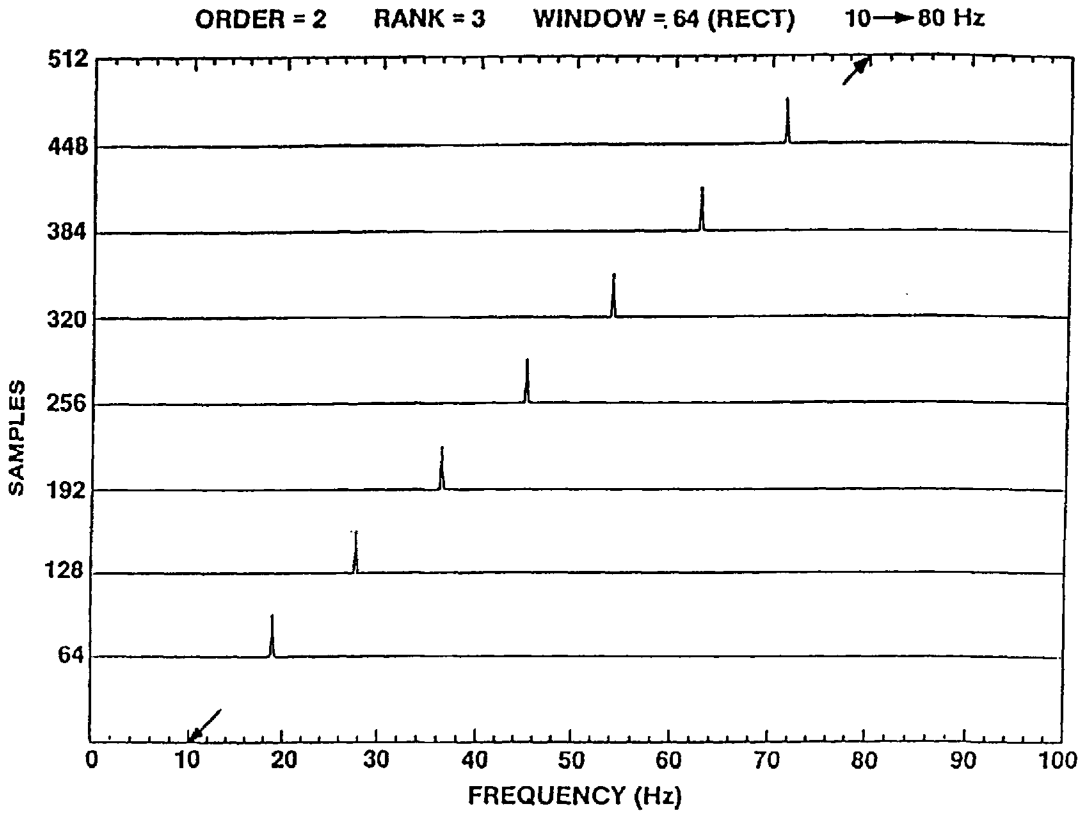

Figure 1: Time-Frequency Representation of a linear FM chirp signal obtained by using an SVD-based Autoregressive Wigner-Ville Distribution (AWVD).

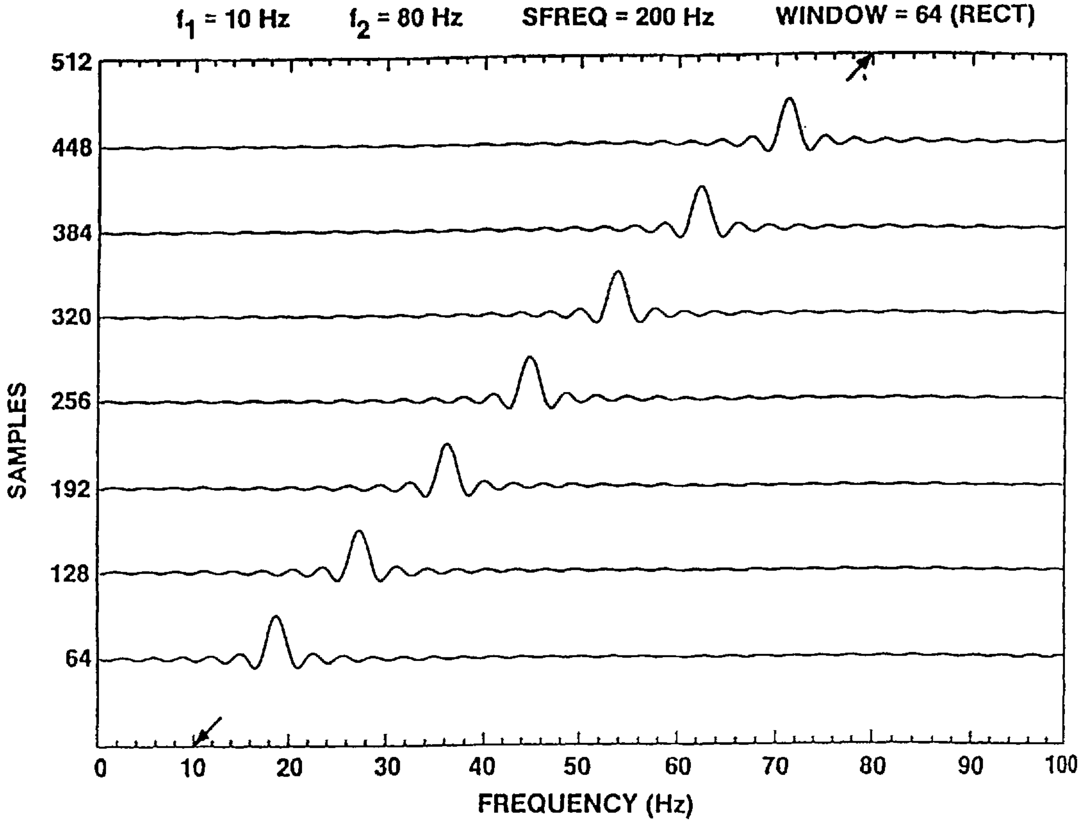

Figure 2: Time-Frequency Representation of a linear FM chirp signal obtained by using the conventional Wigner-Ville Distribution (WVD).

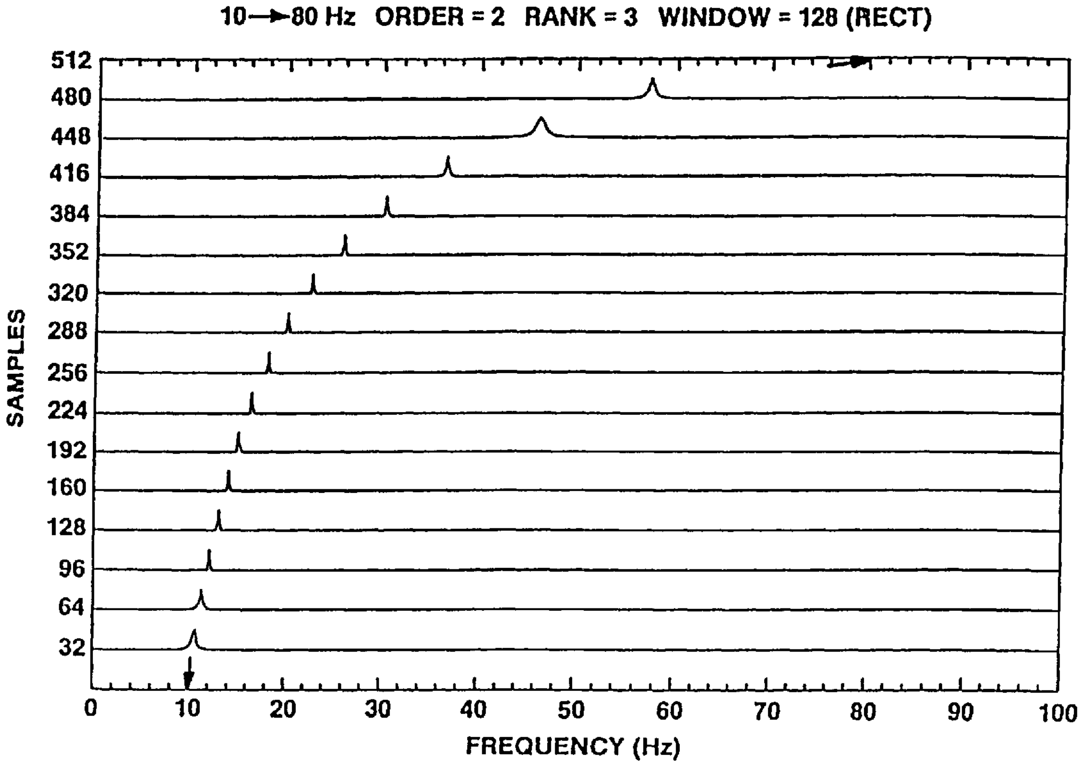

Figure 3: Time-Frequency Representation of a hyperbolic FM chirp signal obtained by using an SVD based Autoregressive Wigner-Ville Distribution (AWVD).

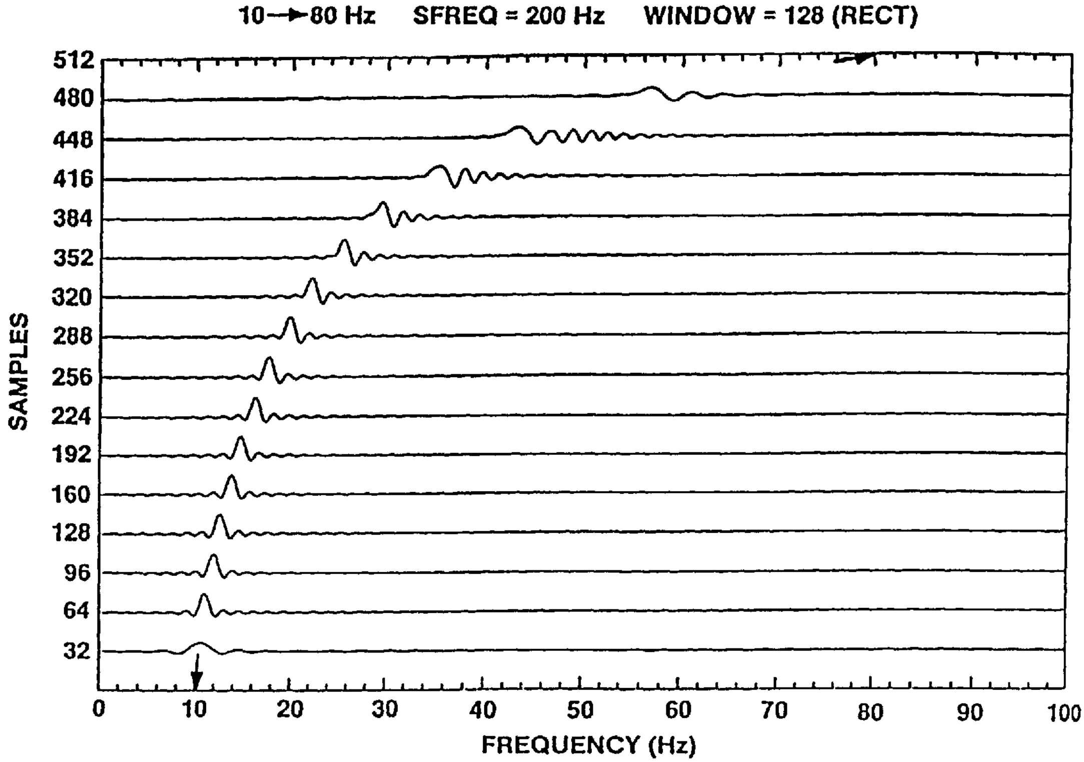

Figure 4: Time-Frequency Representation of a hyperbolic FM chirp signal obtained by using the conventional Wigner-Ville Distribution (WVD).

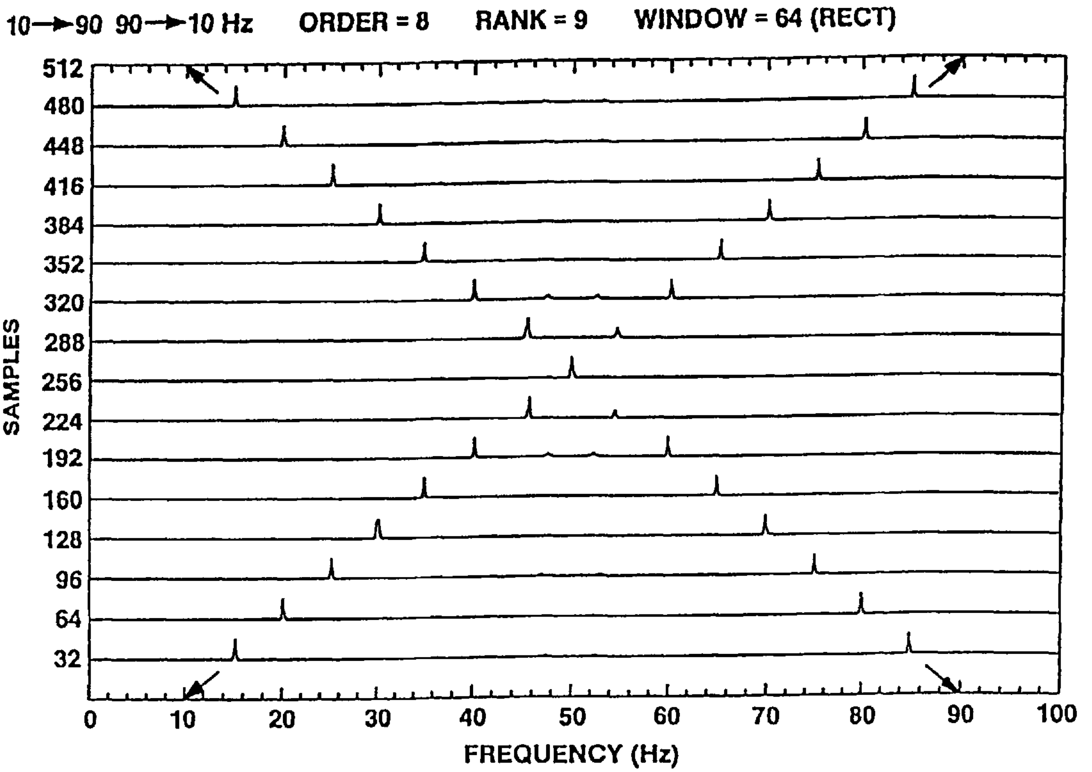

Figure 5: Time-Frequency Representation of two cross-linear FM chirp signals obtained by using an SVD based Autoregressive Wigner-Ville Distribution.

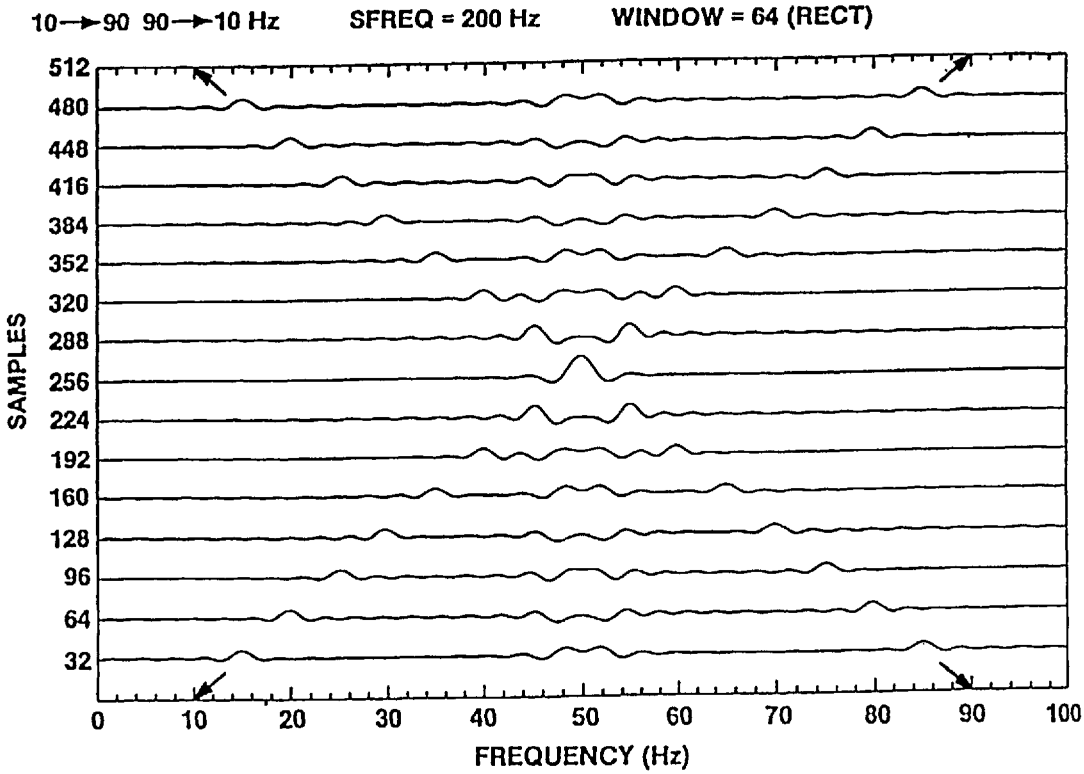

Figure 6: Time-Frequency Representation of two cross-linear FM chirp signals obtained by using the conventional Wigner-Ville Distribution (WVD).